



Building the Research Data Alliance

Dr. Francine Berman

Chair, RDA/US

Hamilton Distinguished Chair in Computer Science

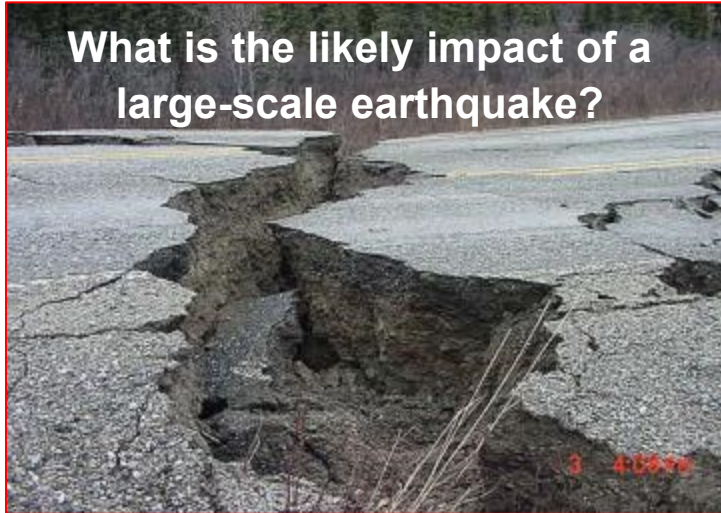
Rensselaer Polytechnic Institute

research data sharing without barriers
rd-alliance.org

Data Driving Solutions to Complex Science and Societal Challenges

2

What is the likely impact of a large-scale earthquake?



Earthquake simulations integrate

Fault geometry data

+ Subsurface data

+ Structure data

+ Population data

+ Physical infrastructure data

Disease spread models couple

Medical data

+ Population data

+ Env't. data

How does disease spread in large urban populations?



How can we increase wheat yields?



Productivity analysis leverages interoperability of

Germplasm data

+ Genetic and phenotypic data

+ Statistical data

+ Bibliographic data

Sharing and Exchange of Data Leading to an Acceleration of Innovation

3

HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

Subscribe: Digital / Home Delivery

Search All NYTimes.com

The New York Times Research

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION | ARTS | STYLE | TRAVEL | JOBS | REAL ESTATE

Do you have at least 4 hours left for sleep?

Search Health 3,000+ Topics

Go

Log in to see what your friends are sharing on nytimes.com. Privacy Policy | What's New

Sharing of Data Leads to Progress on Alzheimer's

By GINA KOLATA
Published: August 12, 2010


What's Popular: Fatally ill, and Making Herself the Lesson

MENU / News & Events / News

BACK TO ALL NEWS

Astronomers Release Unprecedented Data Set on Celestial Objects that Brighten and Dim

BY JONATHAN HANDL



OFFICE OF JUSTICE PROGRAMS

NATIONAL INSTITUTE OF JUSTICE

Research • Development • Evaluation

ABOUT NIJ | WHAT'S NEW

HOME | FUNDING & AWARDS | PUBLICATIONS & MULTIMEDIA | EVENTS | TRAINING | TOPICS

NIJ Home Page > NIJ Journal > NIJ Journal No. 267

In Brief: Expanding Research by Sharing Data

by NIJ staff

NIJ makes data available for future research.



VB

what's hot? | AMAZON | CES 2013 | GOOGLE | 2013 PREDICTIONS | CRUNCHIES AWARDS

THE WALL STREET JOURNAL

THE NEWS YOU WANT. THE PAGES YOU NEED.

Mortgages Fade at Wells Fargo

Will We Wonder What We Did on the Weekend?

Store | Newsletters | Jobs | Events | Research

17 hrs ago: The Moneyball strategy is the future for venture capital firms

2 days ago: Facebook's new regional fan breakdown outs brands that have been buying likes

4 days ago: Disney World's RFID bracelets use data to...

Tech nerds leverage data to help you avoid the flu

InformationWeek Healthcare

Subscribe Today

Software | Security | Cloud | Mobility | Social Business | Big Data | Windows | Global CIO | Government | Healthcare | Education | Financial | SMB | More

Electronic Medical Records | Mobile & Wireless | Clinical Information Systems | Security & Privacy | CPOE | The Patient | Leadership | More Healthcare

HP ProDesk 600 G1 desktop powered by AMD E-Series A6-6400K processor

The power of HP Converged Infrastructure is here. Now you can afford them.

Sharing Psychiatry EHR Data Cuts Readmission Rates

Get InformationWeek Daily

Don't miss each day's hottest technology news, sent directly to your inbox, including occasional breaking news alerts.

research data sharing without barriers
rd-alliance.org

nature
medicine

nature.com > journal home > archive > issue > new

ARTICLE

view full

opt

NATURE MEDICINE | NEWS

日本経済新聞

The delay in sharing research data is costing lives

Josh Sommer

RDA
RESEARCH DATA ALLIANCE

Integrating Data Across Cultures, Scales, Technologies Requires Infrastructure

4

- Common metadata standards
- Interoperability / integration framework
- Data access and preservation policy and practice
- Harmonized standards
- Common economic model for sustaining data
- Digital object identifiers
- Tools for data discoverability
- Etc. ...



Harmonized standards

*Policy and
Practice*



The Research Data Alliance

- Global community-driven organization launched in March 2013 to accelerate data-driven innovation
- Focus is on building the *social, organizational and technical infrastructure* to
 - reduce barriers to data sharing and exchange
 - accelerate the development of coordinated global data infrastructure
- **CREATE → ADOPT → USE:** RDA Working Groups operate for 12-18 months to build and use targeted pieces of data infrastructure



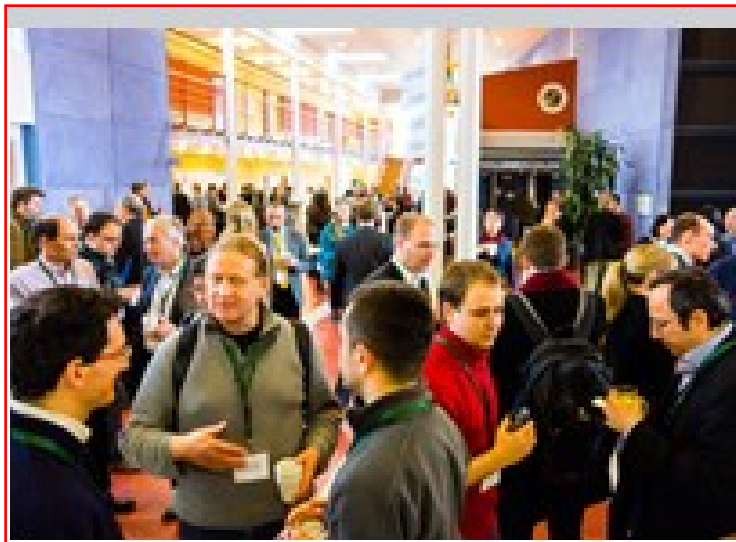
RDA Interest Groups are specific communities mapping out infrastructure that will enable data sharing and exchange, with concrete pieces of infrastructure then proposed as Working Groups

RDA Working Groups are “Tiger Teams” focusing on short-term deliverables:

- **Focused pieces of adopted code, policy, infrastructure, standards, or best practices** that enable data to be shared and exchanged
- **“Harvestable” efforts** for which 12-18 months of work can eliminate a roadblock for a substantial community
- **Efforts that have substantive applicability** to “chunks” of the data community, but may not apply to everyone
- **Efforts for which working scientists and researchers can start today** while more long-term or far-reaching solutions are appropriately discussed in other venues

RDA Launch March 18-20, 2013 Gothenburg, Sweden

- Over 200 participants
- 31 countries
- 5 continents
- > 6,400 tweets
- Public, private, academic sectors
- High-profile Govt. and Science speakers



 **najla**
@najlaoa

Jahanian **#NSF**: the **#RDALaunch** promises to accelerate pace of scientific discoveries. Data is new currency for science

4 DAYS AGO REPLY RETWEET FAVORITE

 **ResearchDataAlliance**
@resdatall

Another theme emerging: RDA as a "neutral" space across domains, organizations, nations, etc. **#RDALaunch**

2 DAYS AGO REPLY RETWEET FAVORITE

 **ResearchDataAlliance**
@resdatall

South Africa can't afford not to participate in RDA **#RDALaunch**

4 DAYS AGO REPLY RETWEET FAVORITE

RDA Interest Groups focusing on infrastructure needed for data sharing and exchange

RDA Interest Groups

- Legal Interoperability (joint with CODATA)
- Engagement
- UPC Code for Data
- Defining Urban Data Exchange for Science
- Marine Data Harmonization
- Repository Audit and Certification
- Preservation e-infrastructure
- Contextual Metadata
- Community Capability Model
- Big Data Analytics
- Agricultural Data, etc.



Agricultural Data Interest Group

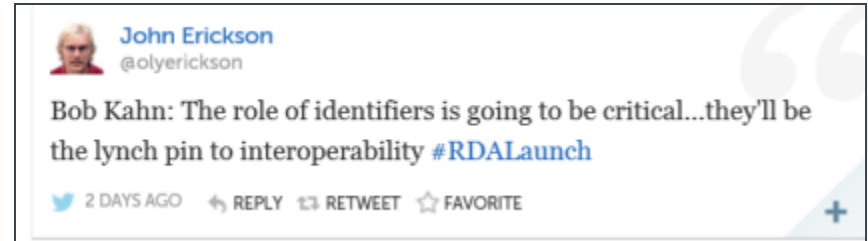
- *Focus:* Share and create interoperability framework for data from agricultural research, in particular wheat
 - *Wheat data combines Germplasm data, genetic and phenotypic data, statistical data, bibliographical data, etc.*
- Group members from India, France, Greece, UN Food and Agricultural Organization



Create-Adopt-Use: RDA Working Groups Creating a Pipeline of Impact-focused Deliverables

Current Working Groups and Groups Under Review

- PID Information Types
- Data Type Registries
- Data Foundation and Terminology (pending)
- Practical Policy (pending)



RDA Working Group: Persistent Identifier (PID) Information Types

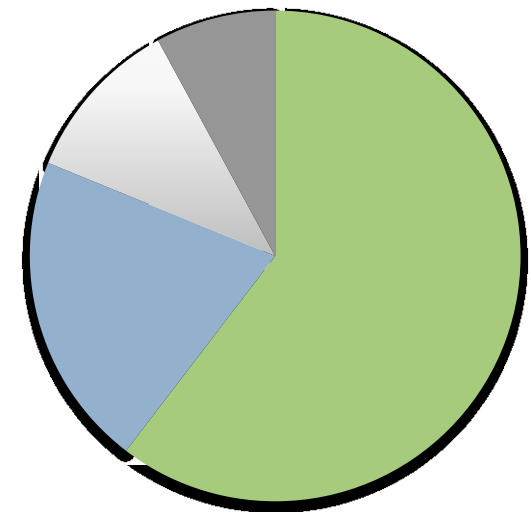
- Group developing a framework for information types associated with unique identifiers; framework will be adopted at DKRZ (German supercomputing Center) and other sites
 - *Agreement about the information associated with PIDs would allow programmers to implement the same API independent of the PID type being used*
 - *Provides a framework for other defined concepts and a prototype API for requesting PID information*

Current Status (as of 4/13): RDA Community = ~700 participants from 44 countries

10

- | | | |
|------------------|---------------|------------------------|
| ■ Albania | ■ Germany | ■ Portugal |
| ■ Australia | ■ Greece | ■ Russia |
| ■ Austria | ■ Iceland | ■ Serbia |
| ■ Bangladesh | ■ India | ■ Singapore |
| ■ Belgium | ■ Iran | ■ South Africa |
| ■ Bulgaria | ■ Ireland | ■ South Korea |
| ■ Brazil | ■ Italy | ■ Spain |
| ■ Canada | ■ Japan | ■ Sweden |
| ■ China | ■ Krygrystan | ■ Switzerland |
| ■ Congo | ■ Kuwait | ■ Taiwan |
| ■ Czech Republic | ■ Netherlands | ■ Turkey |
| ■ Denmark | ■ New Zealand | ■ United Arab Emirates |
| ■ Estonia | ■ Norway | ■ United Kingdom |
| ■ Finland | ■ Palestine | ■ United States (31%) |
| ■ France | ■ Poland | |

RDA by Sector



- Academics (61%)
- Private Sector (21%)
- Public Sector (11%)
- Unknown (8%)

Why is the RDA Important to the US?

11

- **Internationally:** Leading role for US in RDA is strategically important, and increases global competitiveness
- **Nationally:** US has its own data challenges that must be resolved by the RDA/US data community
- *RDA provides a vehicle for implementing the underlying infrastructure required to make new policy approaches work.*



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20503
February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren, JPH
Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

I. Policy Principles

The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

Scientific research supported by the Federal Government drives our economy. The results of that research progress in areas such as health, energy,

Access to digital data sets resulting from research resources and efforts on understanding and data underpins the forecasting industry, as evidenced by many biotechnology innovations, publications and scientific data in digital form, services related to curation, preservation, publications and data for re-use through the impact and accountability of the Federal scientific breakthroughs and innovation, growth and job creation.

The Administration also recognizes that the coordination of peer review, that are essential scholarly publications. It is critical that the important that Federal policy not adversely affect the Federal Government to disseminate

To achieve the Administration's commitment to research and digital scientific data, Federal have clear and coordinated policies for increasing

of digital data, the Obama
Research and Development Initiative." By
rights from large and complex
help solve some the Nation's most



RDA Plenary 2: Opportunity for US Data Leadership

- **RDA Plenary 2: September 16-18 at National Academies in Washington DC**



RDA Plenary is an opportunity

- To convene data-enabled organizations in a “neutral space” that facilitates the implementation and use of global infrastructure and global collaboration
- To highlight US efforts, contributions, and impact with a larger portion of the US data community
- For the **RDA community to progress its work** and expand its efforts



Thank You

enquiries@rd-alliance.org

bermaf@rpi.edu

research data sharing without barriers
rd-alliance.org

Fran Berman